

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

23.06.03

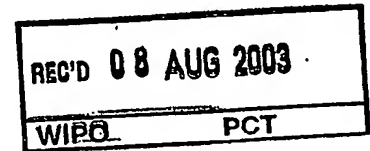
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 2 年    6 月 2 4 日  
Date of Application:

出 願 番 号                      特 願 2 0 0 2 - 1 8 2 5 0 6  
Application Number:  
[ST. 10/C]:                      [ J P 2 0 0 2 - 1 8 2 5 0 6 ]

出      願      人                      松 下 電 器 産 業 株 式 会 社  
Applicant(s):

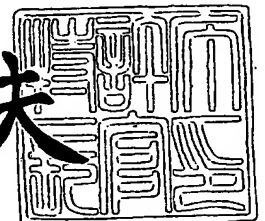


**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2 0 0 3 年    7 月 2 5 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



BEST AVAILABLE COPY

【書類名】 特許願  
【整理番号】 2054041171  
【提出日】 平成14年 6月24日  
【あて先】 特許庁長官殿  
【国際特許分類】 H04N 7/08

## 【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 小林 正明

## 【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 酒井 啓行

## 【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 松井 謙二

## 【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 桑野 裕康

## 【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

## 【代理人】

【識別番号】 100097445

## 【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 メタデータ制作装置及び制作方法

【特許請求の範囲】

【請求項 1】 コンテンツに関連するメタデータの制作装置であって、  
音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、  
上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置。

【請求項 2】 コンテンツに関連するメタデータの制作装置であって、  
音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを具備し、

上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置。

【請求項 3】 上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を備え、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項 1 または 2 のいずれかに記載のメタデータ制作装置。

【請求項 4】 コンテンツに関連するメタデータの制作方法であって、  
音声入力手段と、音声認識手段と、メタデータ制作手段とを用いて、  
上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作方法。

【請求項 5】 コンテンツに関連するメタデータの制作方法であって、  
音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを用いて、

上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付け

て認識し、上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作方法。

【請求項6】 上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を用いて、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項1または2のいずれかに記載のメタデータ制作方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、コンテンツ制作におけるメタデータ制作システム及び方法に関するものである。

【0002】

【従来技術】

近年、映像・音声コンテンツの制作において、これらコンテンツに関連したメタデータの付与することがおこなわれている。

【0003】

しかしながら、上記メタデータの付与は、制作された映像・音声コンテンツのシナリオあるいはナレーション原稿をもとに、制作された映像・音声コンテンツを再生しながらメタデータとすべき情報を確認し、手作業でコンピュータ入力することにより制作する方法が一般的であり、相当な労力の必要な方法であった。

【0004】

【発明が解決しようとする課題】

本願発明は、上記従来の問題点に係る課題を解決することを目的とするものであって、制作された映像・音声コンテンツを再生することによりメタデータとすべき情報を確認し、音声入力でコンピュータに入力することにより制作するシステム及び方法を提供することを目的とする。

【0005】

【課題を解決するための手段】

上記課題を解決するために本願発明は、制作されたコンテンツの再生手段、上

記再生手段で再生された映像信号を表示する映像モニタ手段、上記再生手段で再生された音声信号をモニタする音声モニタ手段、上記映像モニタ手段および上記音声モニタ手段でオペレータが確認した制作すべきメタデータ内容をオペレータの発声によりマイクで収録する音声入力手段、上記音声入力手段により入力された音声信号を認識する音声認識手段、音声認識手段で認識された音声情報をメタデータに変換することによりメタデータを生成するメタデータ生成手段、および上記コンテンツと上記メタデータとを関連づけるため、上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力しタイムコード付メタデータとするタイムコード付与手段とを備えたものである。

#### 【0006】

これにより、従来キーボードで入力し、制作していたメタデータを、音声認識を用いて音声入力し、自動的にタイムコード付きのメタデータを制作することが可能となる。

#### 【0007】

##### 【発明の実施の形態】

本発明の請求項1に係る発明は、コンテンツに関連するメタデータの制作装置であって、音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置である。

#### 【0008】

本発明の請求項2に係る発明は、コンテンツに関連するメタデータの制作装置であって、音声入力手段と音声認識手段とメタデータ制作手段と上記コンテンツに関連した辞書とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて上記コンテンツに関連した辞書に関連付けて認識し、認識したデータを上記メタデータ制作手段によりメタデータに変換することを特徴とするメタデータ制作装置である。

#### 【0009】

本発明の請求項 3 に係る発明は、上記請求項 1 ないし上記請求項 2 に係る発明のメタデータ制作装置であって、上記コンテンツに付与されているタイムコード情報と上記メタデータとを入力し、タイムコード付メタデータを生成するタイムコード付与手段を備え、上記コンテンツと生成された上記メタデータとを関連付けることを特徴とする請求項 1 または 2 のいずれかに記載のメタデータ制作装置である。

#### 【0010】

以下、本発明の実施の形態について図面を用いて説明する。

#### 【0011】

##### (実施の形態 1)

図 1 は、本発明の実施の形態 1 によるメタデータ制作装置の構成を示すブロック図である。図 1 において、1 はコンテンツ再生手段、2 は映像モニタ手段、3 は音声モニタ手段、4 はマイク、5 は音声認識手段、6 はメタデータ生成手段、7 はタイムコード付与手段、8 は辞書である。

コンテンツ再生手段 1 は、たとえば VTR (あるいはハードディスクで構成された映像・音声信号再生装置、あるいは半導体メモリなどのメモリ手段を記録媒体とする映像・音声信号再生装置、あるいは光学記録式または磁気記録式などの回転型ディスクで構成された映像・音声信号再生装置、さらには伝送されてきたあるいは放送されてきた映像・音声信号を再生する映像・音声再生装置などのコンテンツ再生手段) である。上記コンテンツ再生手段 1 は、映像信号出力端子 101、音声信号出力端子 102 およびタイムコード出力端子 103 を具備し、再生された映像信号は端子 101 および 201 を介して、映像モニタ手段 2 に供給され、再生された音声信号は端子 102 および 302 を介して、音声モニタ 3 に供給され、再生されたタイムコードは端子 103 および 703 を介してタイムコード付与手段 7 に供給される。メタデータを制作する制作者 (図示せず) は、映像モニタ手段 2 と音声モニタ手段 3 のいずれかあるいは両方を確認しながら、場合によればシナリオまたはナレーション原稿なども参照しながら、入力すべきメタデータを発声する。マイク 4 は、上記制作者の発声を受付、音声信号に変換して、音声認識手段 5 に供給する。また、必要に応じて、音声認識用の辞書 8 も、音

声認識手段 5 に供給される。音声認識手段 5 で認識された、音声データはメタデータ生成手段 6 に供給され、メタデータあるいはタグに変換される。此のようにして、生成されたメタデータあるいはタグは、コンテンツ自身の内容と時間関係あるいはシーンとの関係を略略一致させるため、タイムコード付与手段 7 にて、コンテンツ再生手段 1 から供給されたタイムコード情報が付与される。

#### 【0012】

より具体的に説明するため料理説明をする場面を一例として想定する。この場合オペレータが、映像モニタ手段 2 の表示画面を確認しながらマイク 4 を通じて“塩 1さじ”と発生すると、音声認識手段 5 にて、辞書 8 を参照しながら、“塩”、“1さじ”と認識されメタデータ生成手段 6 にて各々“塩”、“1さじ”というタグに変換される。なお、音声認識としては、上記音声認識手段 5 に限定されず、一般的に用いられている種々の手段を用いて音声認識し“しお”、“ひとさじ”とのデータを認識してもよい。

なお、一般には、メタデータとは、このようなタグの集合体を意味するものである。タイムコード付与手段 703 では、コンテンツ再生手段 103 からの信号をもとに、タイムコードが付与されたタイムコード付与メタデータが生成される。具体的には、図 2 に示すようなパケットデータが生成される。生成されたメタデータは、そのまま出力されても良く、またハードディスク等の記録媒体に蓄積しても良い。また、上記の実施例においては、パケット形式のメタデータを生成するとして説明したが、特に限定されるものではない。

#### 【0013】

さらに上述した実施例では、コンテンツとしてタイムコードの付与されている動画コンテンツの場合について説明したが、静止画コンテンツあるいは、デジタルデータコンテンツなどの場合には、上記静止画コンテンツあるいはデジタルデータコンテンツを識別するために、動画の場合のタイムコードに相当するコンテンツの番地あるいは番号を用いて上記コンテンツと生成された上記メタデータを関連づけてもよい。

#### 【0014】

なお、一般的には、音声認識には何らかの誤認識が生じる可能性がある。誤認



識が生じた場合、制作されたメタデータ、タグをコンピュータ手段などの情報処理手段を用いて修正することは可能である。

### 【0015】

#### 【発明の効果】

以上説明したように発明は、コンテンツに関連したメタデータの作成あるいはタグ付けを行うに当たり、音声入力による音声認識を用い、且つ、上記メタデータあるいはタグとコンテンツとの時間あるいはシーンとの関連付けを行うため、従来のようなキーボード入力より、効率的に、メタデータの作成やタグ付けを実施することができる。

#### 【図面の簡単な説明】

##### 【図1】

本発明の実施の形態1に係るメタデータ制作装置の構成を示すブロック図

##### 【図2】

本発明の実施形態1に係るタイムコード付きメタデータの一例を示す図

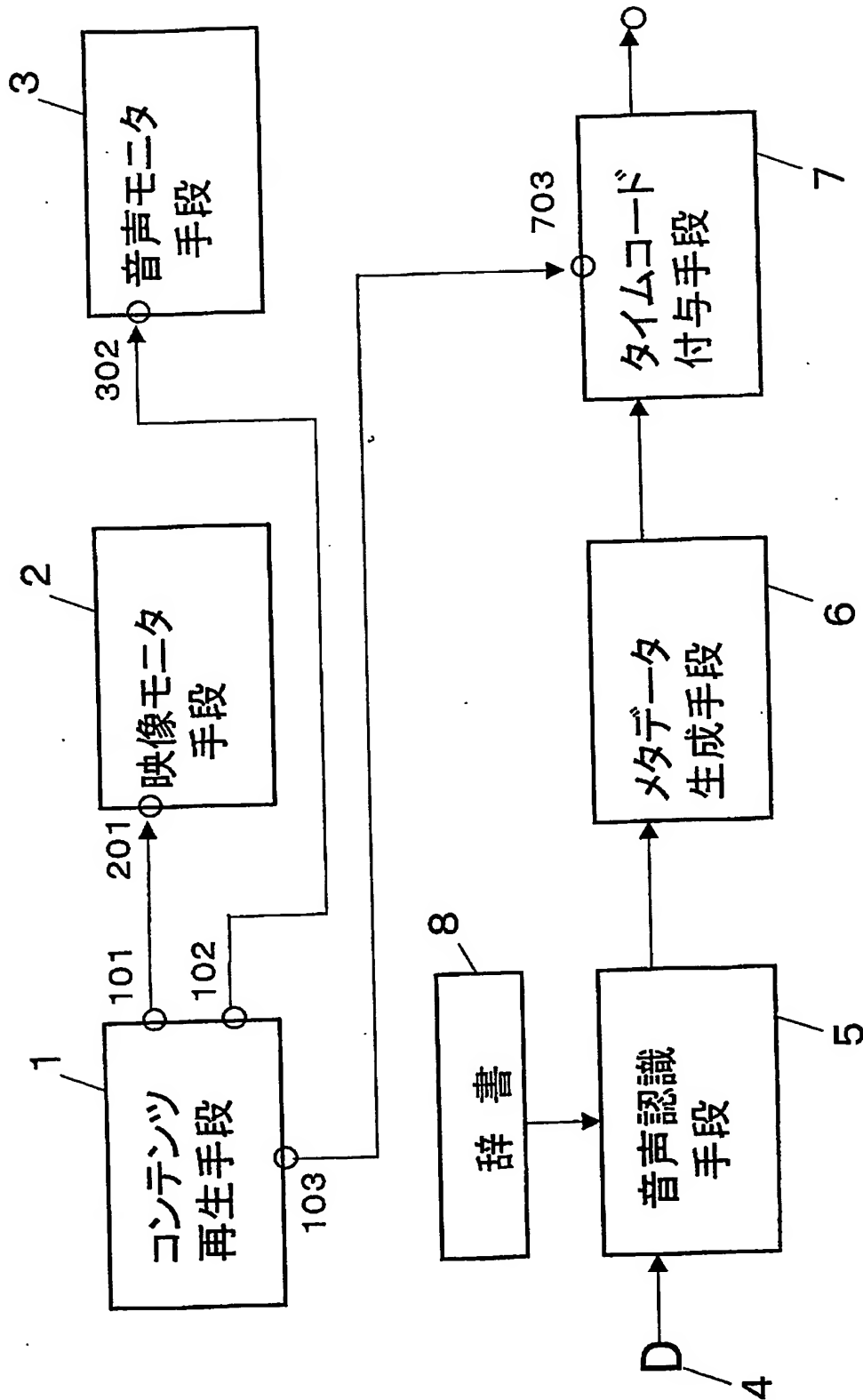
#### 【符号の説明】

- 1 コンテンツ再生手段
- 2 映像モニタ手段
- 3 音声モニタ手段
- 4 マイク
- 5 音声認識手段
- 6 メタデータ生成手段
- 7 タイムコード付与手段
- 8 辞書
- 101 映像出力端子
- 102 音声出力端子
- 103 タイムコード出力端子
- 201 映像入力端子
- 302 音声入力端子
- 703 タイムコード入力端子

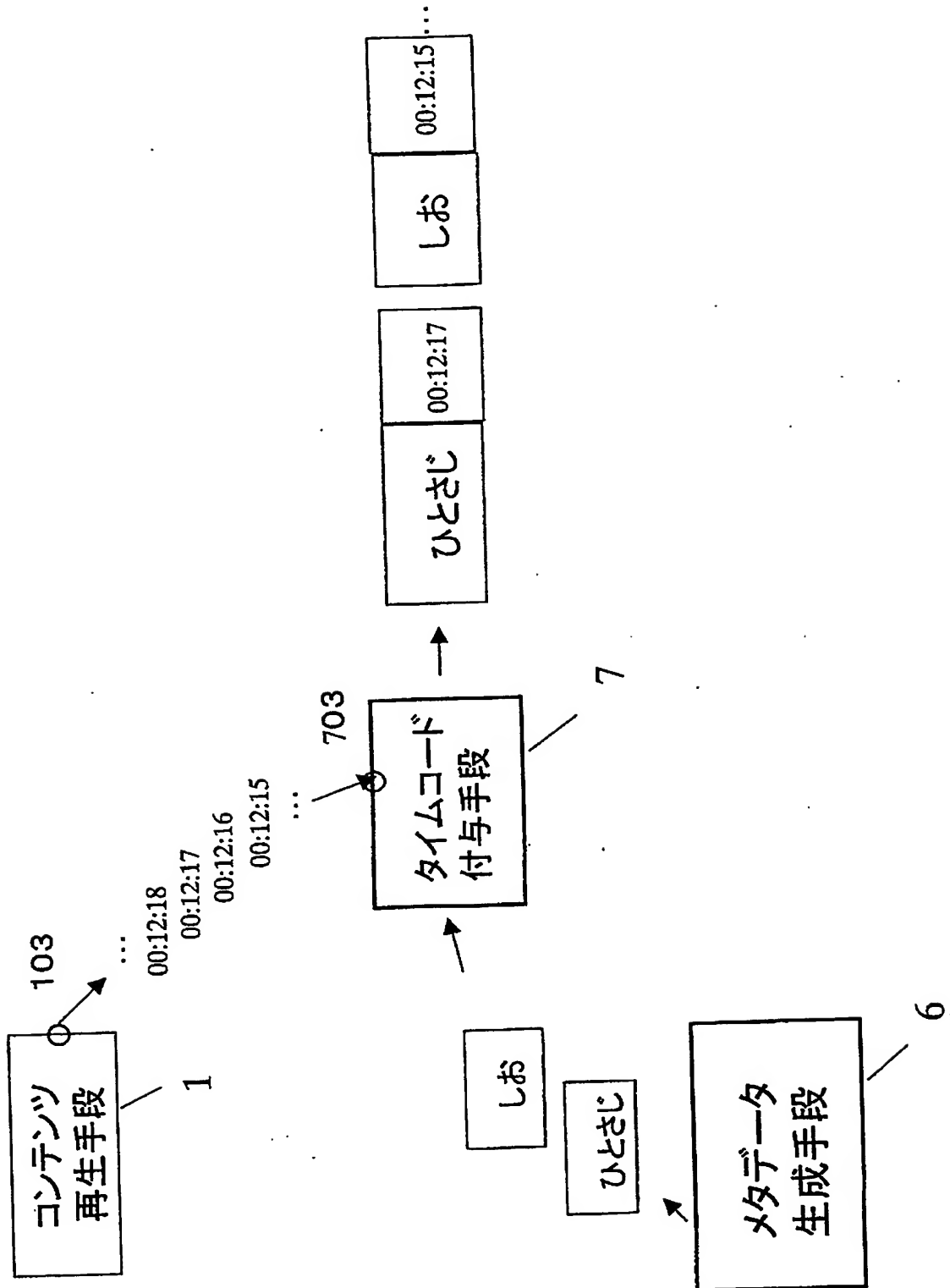
【書類名】

図面

【図 1】



【図 2】



【書類名】 要約書

【要約】

【課題】 映像・音声コンテンツの制作において、これらコンテンツに関連したメタデータの付与するに際し、付与する際の労力を低減することを目的とする。

【解決手段】 音声入力手段と、音声認識手段と、メタデータ制作手段とを具備し、上記コンテンツに関連した情報を上記音声入力手段により入力し、上記入力された音声信号を上記音声認識手段にて認識し、上記メタデータ制作手段によりメタデータに変換することにより、自動的にメタデータ、タグとコンテンツとの時間あるいはシーンとの関連付けを行う。

【選択図】 図1

特願 2002-182506

出 願 人 履 歷 情 報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社